

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 1, 2003, 08:48:11 ; Search time 22.5 Seconds

(without alignments)
5120.944 Million cell updates/sec

Title: US-09-768-781-2

Perfect score: 2543

Sequence: 1 atgaacacaaaccacaaca.....caaggcaaaagtgtgtctga 1389

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09768781/runat.01042003.084729.4929/app.query.fasta_1.1543
-DB=SwissProt.40 -QFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPEGL=0
-DOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09768781 -CGN 1.1 31 @runat.01042003.084729.4929 -NCPUS=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	939	36.9	444	1 XKRY_HUMAN	P51811 homo sapien
2	197.5	7.8	159	1 XKRY_HUMAN	O14609 homo sapien
3	108.5	4.3	497	1 DTPT_ECOLI	O07380 lactobacill
4	108.5	4.3	741	1 YBIO_ECOLI	P75783 escherichia
5	104.5	4.1	521	1 Y725_CABEL	Q10934 caenorhabdi
6	102	4.0	1584	1 BAI1_HUMAN	O14514 homo sapien
7	100.5	4.0	264	1 ATP6_PODAN	P15994 podospora a
8	100.5	4.0	459	1 NUAM_CERSI	O03204 ceratotheri
9	99.5	3.9	551	1 YNBI_YEAST	P53730 saccharomyc
10	99	3.9	233	1 ATP1_OCHNE	Q40607 ochrosphaer
11	99	3.9	458	1 ACHO_HUMAN	Q05901 homo sapien
12	98	3.9	805	1 VPH1_SCHPO	O13742 schizosacch
13	97.5	3.8	325	1 MCSR_RAT	P35345 rattus norv
14	97	3.8	382	1 NUZM_CHLRE	P08740 chlamydomon
15	97	3.8	527	1 NUSM_CAEEL	P24896 caenorhabdi
16	96.5	3.8	384	1 NUSM_ANOAR	P51899 anopheles a
17	96.5	3.8	580	1 NUSM_ANOGA	P34854 anopheles g
18	96	3.8	371	1 NYGR_MOUSE	Q61212 mus musculus

19	96	3.8	1033	1 YD56_SCHPO	Q10309 schizosacch
20	95.5	3.8	252	1 ATP6_NEUCR	P37212 neurospora
21	95.5	3.8	255	1 ATP6_YARLI	Q36258 yarrowia li
22	95.5	3.8	325	1 MCSR_MOUSE	P41149 mus musculu
23	95.5	3.8	405	1 YMP0_YEAST	Q03687 saccharomyc
24	95.5	3.8	429	1 SECY_AQUAE	O66491 aquifex aeo
25	95	3.7	469	1 NUON_BUCAI	P57264 buchnera ap
26	94.5	3.7	438	1 MNT2_PSEAE	Q9RPF2 pseudomonas
27	94.5	3.7	449	1 G64F_DROME	P83297 drosophila
28	94.5	3.7	468	1 CBIA_FUGRU	Q98894 fugu rubrip
29	94.5	3.7	786	1 APE1_SULTO	O96429 sulfolobus
30	94.5	3.7	865	1 CYAA_PROMI	O59685 proteus mir
31	94	3.7	592	1 FZD1_CHICK	O57328 gallus gall
32	94	3.7	607	1 NUSM_MOUSE	P03921 mus musculu
33	94	3.7	692	1 NUSC_LARPA	P06284 marchantia
34	94	3.7	715	1 LCCL_LACLA	Q9C3B8 lactococcus
35	93.5	3.7	456	1 G64A_DROME	P83293 drosophila
36	93.5	3.7	971	1 CSE1_MOUSE	O9ERK4 mus musculu
37	93.5	3.7	1564	1 PDRA_YEAST	P51533 saccharomyc
38	93	3.7	406	1 HOF2_HAEIN	P44821 haemophilus
39	93	3.7	497	1 NUZM_CHOCR	P48903 chondrus cr
40	93	3.7	871	1 YAYD_SCHPO	Q09766 schizosacch
41	92.5	3.6	327	1 GPS2_HUMAN	Q13227 homo sapien
42	92	3.6	248	1 ATP1_WARPO	P06289 marchantia
43	92	3.6	437	1 NUAM_ALBCO	P48914 albinaria c
44	92	3.6	455	1 ACHO_CHICK	P43679 gallus gall
45	92	3.6	554	1 NUSM_APLLI	P34855 apis mellif

ALIGNMENTS

RESULT 1

XX HUMAN	ID	XX HUMAN	STANDARD;	PRT;	444 AA.
AC	P51811;				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Membrane transport protein XK (Kx antigen).				
GN	XK.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94273191; PubMed=8004674;				
RA	Ho M., Chelly J., Carter N., Danek A., Crocker P., Monaco A.P.;				
RT	"Isolation of the gene for McLeod syndrome that encodes a novel				
RT	membrane transport protein.";				
RL	Cell 77:869-880(1994).				
RN	[2]				
RP	REVIEWS TO 204-205.				
RA	Ho M.;				
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.				
CC	-!- FUNCTION: MAY BE INVOLVED IN SODIUM-DEPENDENT TRANSPORT OF NEUTRAL				
CC	AMINO ACIDS OR OLIGOPEPTIDES.				
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-!- TISSUE SPECIFICITY: HIGH LEVELS IN SKELETAL MUSCLE, HEART, BRAIN,				
CC	AND PANCREAS; LOW LEVELS IN PLACENTA, LUNG, LIVER, AND KIDNEY.				
CC	-!- POLYMORPHISM: XK IS RESPONSIBLE FOR THE KX BLOOD GROUP SYSTEM.				
CC	-!- DISEASE: DEFECTS IN XK ARE THE CAUSE OF MCLEOD SYNDROME, AN X-				
CC	LINKED MULTISYSTEM DISORDER CHARACTERIZED BY ABNORMALITIES IN THE				
CC	NEUROMUSCULAR AND HEMATOPOIETIC SYSTEMS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				

DR EMBL; Z32684; CAA83632.2; --
 DR Genew; HGNC:12811; XK.
 DR MIM; 314850; --
 KW Transmembrane; Transport; Amino-acid transport; Blood group antigen.
 FT DOMAIN 1 2
 FT TRANSMEM 3 23 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 24 37 POTENTIAL.
 FT TRANSMEM 38 58 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 59 68 POTENTIAL.
 FT TRANSMEM 69 89 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 90 140 POTENTIAL.
 FT TRANSMEM 141 161 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 162 171 POTENTIAL.
 FT TRANSMEM 172 192 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 193 208 POTENTIAL.
 FT TRANSMEM 209 229 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 230 235 POTENTIAL.
 FT TRANSMEM 236 256 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 257 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 298 POTENTIAL.
 FT TRANSMEM 299 317 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 339 349 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 350 370 POTENTIAL.
 FT TRANSMEM 371 444 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 444 AA; 50902 MW; 6F90B0B45659AIDA CRC64;

Alignment Scores:
 Pred. No.: 1,16e-69 Length: 444
 Score: 939,00 Matches: 180
 Percent Similarity: 64.68% Conservative: 80
 Best Local Similarity: 44.78% Mismatches: 132
 Query Match: 36.92% Indels: 10
 DB: 1 Gaps: 4

US-09-768-781-2 (1-1389) x XK_HUMAN (1-444)

QY 136 TTTCATTAGTACCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTG 195
 Db 3 PheProAlaSerValLeuAlaSerValPheLeuPheValAlaGluThrThrAlaAlaLeu 22
 QY 196 TACATGTTAGATCTATCGAAGAATAGTGAACTTACTGTGATGACATACACCTTTTCT 255
 Db 23 SerLeuSerSerThrThrArgSerGlyGlyAspArgMetTrpGlnAlaLeuThrLeuLeu 42
 QY 256 TTCTTTATGTTTTCATCCATTATGTCAGTGTACCTCATCTTTTGTGCCAGAGATCTA 315
 Db 43 PheSerLeuLeuProCysAlaLeuValGlnLeuThrLeuLeuPheValHisArgAspLeu 62
 QY 316 GCCAAGATAAACGGCTATCATTTATTTATGATCTAATCTCTTGGGACCTGTTATCAGA 375
 Db 63 SerArgAspArgProLeuValLeuLeuHisLeuLeuGlnLeuGlyProLeuPheArg 82
 QY 376 TGTGTGGAGCCATGATTAAGTACTCTCACACTGTGGGAAGAAGAGGAGGAGGAGGCC 435
 Db 83 CysPheGluValPheCysIleTyr-----PheGlnSerGlyAsnAsnGluGluPro 99
 QY 436 TATGTGAGCTTACCCGAAAGAAG--ATGCTATATAGATGCGGAGGAGGTCTCATAGAA 492
 Db 100 TyrValSerIleThrLysArgGlnMetProLysAsnGlyLeuSerGluGluLeuGlu 119
 QY 493 TGGAGGTGGGCCTCCATCCGGACCTCGCTATGACCGCAATGCTCAACACGATG 552
 Db 120 LysGluValGlyGlnAlaGluGlyLysLeuIleThrHisArgSerAlaPheSerArgAla 139
 QY 553 TCACAGATCAAGCTTCTGGGCTAGTCCCGAGTACCTATGACCTATGACCTATGTCAGC 612
 Db 140 SerValIleGlnAlaPheLeuGlySerAlaProGlnLeuThrLeuGlnLeuThrIleSer 159
 QY 613 CTGATCTCTCAGAGGTTCCCTGGGTAGAGTGTGTCTAATGGTATTTTCCCTGGTATCT 672
 Db 160 ValMetGlnGlnAspValThrValGlyArgSerLeuLeuMetThrIleSerLeuLeuSer 179

QY 673 GTCACCTATGGGCCACCCCTTTTGAATATGTTGGTATCCAGATCAAGTACGATGACTAC 732
 Db 180 IleValTyrGlyAlaLeuArgCysAsnIleLeuAlaIleLysIleLysAspGluTyr 199
 QY 733 AAGATTCCCTTGGGCCACTAGAAAGTCTCTGTCATCACCATCTGGCGGACATTTGGAGATC 792
 Db 200 GluValLysValLysProLeuAlaTyrValCysIlePheLeuTrpArgSerPheGluIle 219
 QY 793 ACTTCCCGCTCTCGATTCTGGTCTCTTCTCAGCCACTTTGAAATTTGAAGCTCTGCCC 852
 Db 220 AlaThrArgValValValLeuValLeuPheThrSerValLeuLysThrTrpValValVal 239
 QY 853 TTCTAGTGTCAACTTCTCTGATCATCTCTTGTAGCCCTCGATTAACTTCTGGAGAGT 912
 Db 240 IleLeuIleAsnPhePheSerPhePheLeuTyrProTrpIleLeuPheTrpCysSer 259
 QY 913 GGTGCCAGATGCCCAATAACATTGAGAAAAAATTCAGCCGGGTCCGACACTCTGTGGTTC 972
 Db 260 GlySerProPheProGluAsnIleGluLysAlaLeuSerArgValGlyThrThrIleVal 279
 QY 973 CTGATTTCACTACCATCTCTATGCTGGCATCAACTTCTTCTGTGCTCAGCTTTGCGAG 1032
 Db 280 LeuCysPheLeuThrLeuLeuTyrThrGlyIleAsnMetPheCysTrpSerAlaValGln 299
 QY 1033 TTGAGGTGGCAGACAGATCTCTCGACAAAGGGCAGAACTGGGGACATATGGGCTG 1092
 Db 300 LeuLysIleAspSerProAspLeuIleSerLysSerHisAsnTrpTyrGlnLeuLeuVal 319
 QY 1093 CACTATAGTGTGAGTGGTAGAGAATGTGATCATGTGCTTGGTGTAAAGTCTTTTGA 1152
 Db 320 TyrTrpMetIleArgPheIleGluAsnAlaIleLeuLeuLeuLeuTrpTyrLeuPheLys 339
 QY 1153 GTGAAAGTGTACTGAATTAATCTGTCATTCTCTTGTGCTTGCAGCTCATTTATCTTAT 1212
 Db 340 ThrAspIleTyrMetTyrValCysAlaProLeuLeuValLeuGlnLeuLeuIleGlyTyr 359
 QY 1213 CTGATTTCCATTGGCTCATGCTCTCTTTTCCAGTACTTGCATCCATTCGGGCTCCTC 1272
 Db 360 CysThrAlaIleLeuPheMetLeuValPheTyrGlnPheHisProCysLysLysLeu 379
 QY 1273 TTCACCCATATAGTAGTAGAC-----TACCTCCATTGTGTCTGC-----TGT 1314
 Db 380 PheSerSerSerValSerGluGlyPheGlnArgTrpLeuArgCysPheCysTrpAlaCys 399
 QY 1315 CACCAG 1320
 Db 400 ArgGln 401

RESULT 2
 XKRY_HUMAN STANDARD; PRT; 159 AA.
 ID XKRY_HUMAN
 AC O14609;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Testis-specific XK related protein Y.
 GN XKRY.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98022381; PubMed=9381176;
 RA Lahn B.T., Page D.C.;
 RT "Functional coherence of the human Y chromosome.";
 RL Science 278:675-680(1997).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: TESTIS-SPECIFIC.
 CC -!- SIMILARITY: SOME TO MEMBRANE TRANSPORT PROTEIN XK.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)

DR EMBL; AF000997; AAC51844.1; -
 DR Genbank; HGNC:18571; XKRY.
 DR MIM; 400015; -
 KW Transmembrane.
 FT TRANSMEM 1 21 POTENTIAL.
 FT TRANSMEM 45 65 POTENTIAL.
 FT TRANSMEM 72 92 POTENTIAL.
 SQ SEQUENCE 159 AA; 18083 MW; B9B81DD842DEFOAB CRC64;

Alignment Scores:
 Pred. No.: 1,236-08 Length: 159
 Score: 197.50 Matches: 52
 Percent Similarity: 58.5% Conservative: 37
 Best Local Similarity: 34.2% Mismatches: 50
 Query Match: 7.7% Indels: 13
 DB: 1 Gaps: 3

US-09-768-781-2 (1-1389) x XKRY_HUMAN (1-159)

QY 642 AGTGTGCTAAGTATTTTCCTGCTATCTGTCACCTATGGGGCCACCTTTGCAATAT 701
 Db 6 SerileAlaAspIlePheProLeuIleSer-CysValGlyAlaIleHisCysAsnII 25
 QY 702 GTTGGCTATCCAGATCAAGTACGATGACTACAGATTCGCTTGGGCCACTAGAACTCCT 761
 Db 25 eLeuAlaIleArgThrGlyAsnAspPheAlaIleLeuLeuGlnValIleLysLeuII 45
 QY 762 CTGCATCACCATCTCGCGGACATTTGAGATCACTTCGCGCTCCTGATTCGTGCTCTT 821
 Db 45 eTyxLeuMetIleTrpHisSerLeuValIleLeuSerProValValThrLeuAlaPhePh 65
 QY 822 CTCAGCCATTTGAATTCAGGCTGTGCGCTTCCTCTAGTCTCAACTTCTGTGATCATCT 881
 Db 65 eProAlaSerLeuValGlnGlySerLeuHisPheLeuLeuIleIleTyxPheValLeuLe 85
 QY 882 CTTTGAGCCCTGGATTAAGTTCCTGGAGAAGTGTGGCCAGATGCCCAATAACATTGAGAA 941
 Db 85 uLeuThrProTrpLeuGluPheSerLysSerGlyThrHisLeuProSerAsn-ThrLysI 105
 QY 942 AAATTCAGCCGGTGGCGCACTCTGTGG-----TCCTGATTCAGTCCACAT 989
 Db 105 leIleProAla-----TyrTrpValSerMetAspAla-TyrLeuAsnHisAla 120
 QY 990 CTTCTATGCTGCATCAACTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049
 Db 121 SerIleCysHisGln-PheSerCysLeuSerAlaValLysLeuGlnLeuSerAsnGI 140
 QY 1050 AGATCTCGCGCAAGGGCAGAACTGG 1077
 Db 140 uGluLeuIle---ArgAspThrArgTrp 148

RESULT 3

ID DTPT LACHE STANDARD; PRT; 497 AA.
 AC 007380;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE DI-/tripeptide transporter.
 GN DTPT.
 OS Lactobacillus helveticus.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1587;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCDO 2712;

RX MEDLINE=97316430; PubMed=9172341;
 RA Nakajima H., Hagting A., Kunji E.R.S., Poolman B., Konings W.N.;
 RT "Cloning and functional expression in *Escherichia coli* of the gene
 encoding the di- and tripeptide transport protein of *Lactobacillus*
 RT *helveticus*.";
 RL Appl. Environ. Microbiol. 63:2213-2217 (1997).
 CC -|- FUNCTION: PROTON-DEPENDENT UPTAKE OF DI- OR TRI-PEPTIDES.
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)

CC EMBL; U77486; AAC45382.1; -
 DR InterPro; IPR000109; PTR2.
 DR Pfam; PF00854; PTR2; 1.
 DR TIGRfam; TIGR00923; 2A1701; 1.
 DR PROSITE; PS01022; PTR2_1; 1.
 DR PROSITE; PS01023; PTR2_2; 1.
 KW Peptide transport; Transport; Transmembrane.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 26 46 POTENTIAL.
 FT TRANSMEM 57 77 POTENTIAL.
 FT TRANSMEM 84 104 POTENTIAL.
 FT TRANSMEM 119 139 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 227 247 POTENTIAL.
 FT TRANSMEM 294 314 POTENTIAL.
 FT TRANSMEM 321 341 POTENTIAL.
 FT TRANSMEM 372 392 POTENTIAL.
 FT TRANSMEM 452 472 POTENTIAL.
 SQ SEQUENCE 497 AA; 55469 MW; 69C064FF9A0A26A5 CRC64;

Alignment Scores:

Pred. No.: 0,324 Length: 497
 Score: 108.50 Matches: 85
 Percent Similarity: 36.76% Conservative: 65
 Best Local Similarity: 20.83% Mismatches: 133
 Query Match: 4.27% Indels: 125
 DB: 1 Gaps: 18

US-09-768-781-2 (1-1389) x DTPT_LACHE (1-497)

QY 310 GATCTAGCCAAAGATAACCGCTATCATTTATTCATCTCTTGGGACCTGTT 369
 Db 110 AspProArgAspAlaGlyPheSerIlePheValPheGlyLeuAsnLeuGlySerIle 129
 QY 370 ATCAGA----- 375
 Db 130 IleAlaProTrpLeuValProTrpAlaAlaGlnGlyPheGlyValHisIlePheGlySer 149
 QY 376 -----TGTGGAGGCCATGATT 393
 Db 150 GlnLeuAsnPheHisAlaGlyPheSerLeuAlaAlaValGlyMetPhePheGlyLeuVal 169
 QY 394 AGTACCTCACACTGTGGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 453
 Db 170 GlnTyxValLeuGlyGlyLysLysTyxLeuSerThrGlu-----SerLeuThrPro 186
 QY 454 AGAAGATGCTAATAGATGCGGAGGAGGTG---CTGATAGATGGGAGGTGGGCACTCC 510
 Db 187 AsnAspProIleAspLysGlyAspLeuLeuAsnValIleLysTrpValIleIleIle 206
 QY 511 ATCCGGACCTGGCTATGCACCCCAATGCCACAAAGTATGTCACAGATCCCAAGCCTTC 570
 Db 207 IleAlaIleValAla-----IleLeuAlaAla 215

```
QY 571 CTGGGCTCAGTGGCCCGAGCTGACCTATCAGCTCTATGTGAGCCCTG-----ATCTCT 621
Db 216 MetAlaGlyValGlyGlnLeuSerValAspAsnValIleThrLeuLeuThrIleLeuAla 235
QY 622 GCAGAGTTCCTCCCTGGGTAGAGTTGCTGAATGTAATTTTCC-----CTGGTATCTGTC 675
Db 236 IleAlaLeuProIleTyrTyrPheValMetMetPheArgSerSerLysValThrLysIle 255
QY 676 ACCTATGGGGCCACCTTTTGCATATG-----702
Db 256 GluLeuGlyIleHisLeuLeuProValSerLeuLysAsnArgLeuPhePheLysGly 275
QY 703 -----TTGCTATCCAGATCAAGTACGACTACAAG 735
Db 276 TyrLysArgLeuLysGlnIleGlnLeuAlaIleLysArgGlnSerPheIle 295
QY 736 ATTGGCTTGGGCGACTAGAGTCTCTGATCACCATCTGGCGGACATTTGGAGTCACT 795
Db 296 IleLeuIleAlaLeuIleIleMetAlaSerIleLeuIleProAsnLysValIleAla 315
QY 796 TCCCGCTCTCTGATCTGCTGCTCTTC---TCAGCCACTTGAATGAAGGCTGTGCC 852
Db 316 LysHisLeuLeuLysValLeuValPheTyrTrpIleGlyLeuAsnLeuIlePro 335
QY 853 TTC-----CTAGTGCTCAACTCTCTGATCATCTCTTTGAGCCCTGATTAAG----- 900
Db 336 PheSerThrPheValLeuSerPhe-----LeuPheLeuAspTyrIleLysHisMet 352
QY 901 TTCTGGAGAGTGGTGGCCAGATGCCCAATAACATTGAGAAAC----- 945
Db 353 PheLysLysGluGlyGlnAlaLysThrLysGlyLysSerArgIleHisGly 372
QY 946 -----TTCAGCCCG-----GTCCGCACTCTGGTGGTCTG 975
Db 373 IleGluIleProLeuPheLeuArgGlnLeuIleAsnIlePheThrLeuIleIleLeu 392
QY 976 ATTTAGTCACTACCTCTC---TATGCTGGCATCACTTCTCTTCG----- 1017
Db 393 GluGlyThrLeuPheAspGluAsnGlyValGluValAsnIleAlaGluHisProVal 412
QY 1018 -----TGCTCAGCTTTGCGAGTTGAGTTGGCGACAGAGATCTCTCGACAAAGGCGAG 1071
Db 413 GlnGlyTyrThrGluLeuAsnIleAsnLeuAsnLysAspSerIleAspLeuTrpAla 432
QY 1072 AACTGGGACATATGGCGCTGCACATATAGTGTGAGGTGGTAGAATGTGATCGTC 1131
Db 433 AspTrp-----IleGlnSerVal----- 438
QY 1132 TTGGTTTAAAGTTCTTGGAGTGAAGTCTTACTGAATCTACTCTCTTCATTGCCC 1191
Db 439 -----AlaLysTyrLeuLeuAsnIleMetTyrThrAlaAspVal 451
QY 1192 TTGAGCTCATTATTGTTTATTCATTTCCATTTGGCTTCATCTCTTCTTCCAGTAC 1251
Db 452 IleValIleIleIlePheTyrLeuValLysMetAlaAlaLeuTrpTrpAlaTrpSerTyr 471
QY 1252 TTGCATCCATTTGGCTCACTCTTC 1275
Db 472 Ile---ProLeuSerThrValPhe 478

RESULT 4
YBIO_ECOLI
ID YBIO_ECOLI STANDARD; PRT; 741 AA.
AC P75783; Q9R7S0; Q9R7S2; Q9ZBC5;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ybio.
GN YBIO OR B0808.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
```

```
RN
SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN
SEQUENCE FROM N.A.
RC STRAIN=K12;
MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
Kimura S., Kitagawa K., Makino K., Masuda S., Miki T., Mizobuchi K.,
Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
Sampai G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
Yano M., Horiuchi T.;
RA "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155 (1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0003 FAMILY.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
frameshift in position 77.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; AE000183; AAC73895.1; ALT_INIT.
DR EMBL; D90717; BAA35474.1; -.
DR EMBL; D90718; BAA35479.1; ALT_FRAME.
DR EMBL; D90718; BAA35480.1; ALT_FRAME.
DR EcoGene; EGI3320; Ybio.
DR InterPro; IPR001880; MSion_channel.
DR Pfam; PF00924; MS_channel_1.
DR PROSITE; PS01246; UPF0003; FALSE_NEG.
DR KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 225 245 POTENTIAL.
FT TRANSMEM 268 288 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 343 363 POTENTIAL.
FT TRANSMEM 374 394 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 466 486 POTENTIAL.
FT TRANSMEM 509 529 POTENTIAL.
FT TRANSMEM 533 553 POTENTIAL.
FT TRANSMEM 562 582 POTENTIAL.
FT TRANSMEM 608 628 POTENTIAL.
FT TRANSMEM 662 682 POTENTIAL.
FT DOMAIN 125 30 POLY-THR.
FT DOMAIN 195 202 POLY-LEU.
SQ SEQUENCE 741 AA; 81903 MW; 69B0A27678976DB5 CRC64;

Alignment Scores:
Pred. No.: 0.347 Length: 741
Score: 108.50 Matches: 97
Percent Similarity: 37.88% Conservative: 92
Best Local Similarity: 19.44% Mismatches: 177
Query Match: 4.27% Indels: 133
DB: 1 Gaps: 25

US-09-768-781-2 (1-1389) x YBIO_ECOLI (1-741)
```

```
QY 22 TCAGAAAGAACCTCGACAAATGGACAGAGTTTATGAATTCCTGAGGAGCCAAATGTGGAT 81
Db 23 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 ThrSerArgLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 79
QY 82 CCGTTTCTATCTCTGGAGGAAGATGTCCTGGTGGAGCCAAACCCCGATTTACTTTTCCA 141
Db 83 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 ProVal-----ProLysIleValProPro 87
QY 142 TTTAGCATCTTTTCTCCACCTTTTG-----TACTGTGG 177
Db 88 ThrLeuValGluGluClnThrValLeuGlnLysValThrGluValSerArgHisTyrGly 107
QY 178 GAGGTGTCATCTGCTTTGTACATGGTTAGAAATCTATCGAAG-----219
Db 108 GluAlaLeuSerAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 126
QY 220 -----AATAGTGAACCTACTGCGATGACATACACC-----249
Db 127 LysProPheAsnProGlnThrPheSerAsnAlaLeuThrHisPheSerMetLeuAlaVal 146
QY 250 -----TTTTCTTCTTTATGTTTTCATCATTTATGTCAGTGTGACCTCATTTTGTG 303
Db 147 LeuValPheGlyPheTyrTrp-----LeuIleArgLeuCysAlaLeuProLeu 162
QY 304 CACAGAGATCTAGCC-----AAGATAAACCGCTATCATTTATTATGCAT 348
Db 163 TyrArgLysMetGlyClnTrpAlaArgGlnLysAsnArgLysSerAsnTrpLeuGln 182
QY 349 CTA-----ATCCTCTTGGGACCTGTTATCAGATGTTGGAGGCCATGATTAAAGTACCTC 402
Db 183 LeuProAlaMetIleIleGlyAlaPheIle-----IleAspLeuLeuLeuAlaLeu 200
QY 403 ACACGTGTGAAGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 462
Db 201 ThrLeuPheValGlyGln-----206
QY 463 CTAATAGATGGCGAGGAGGTGCTGATAGAAATGGAGGTGGCGACCTCCATCCGAGCCCTG 522
Db 207 -----ValLeuSerAspAsnLeuAsnAlaGlySerArgTrpIle 219
QY 523 GCTATGCACCGCAATGCTTACAAACGTATGTACAGATCCAGACCTTCTCGGCTCAGTG 582
Db 220 AlaPheGlnSerLeuPheLeuAsnAlaPheAlaLeuIleGluPhePheLysAlaVal 239
QY 583 CCCAGCTGACCTAT-----CAGCTCTATGTGAGCTGAGCTGATCTTCGAGAG 627
Db 240 LeuArgLeuIlePheCysProAsnValAlaGluLeuArgProPheThrIleGlnAspGlu 259
QY 628 GTTCCC-----CTGGGTAGAGTTGTCTAATGGTATTTTCCCTGTATCTGTACCTAT 681
Db 260 SerAlaArgTyrTrpSerArgLysLeuSerTrpLeuSerSerLeuIleGly-----Tyr 277
QY 682 GGGGCGACCTTTGCAATATGTTGGTATCCAGATCAAGTACAGTACACTACAAGATTGCG 741
Db 278 GlyLeu-----IleValAlaValProIleIleSerAsnGlnValAsnValGln 293
QY 742 TTGGGGCCATAGAA-----GTCTCTGTGCATCACCATCTGG-----777
Db 294 IleGlyAlaLeuAlaAsnValIleIleMetLeuCysMetThrValTrpAlaLeuTyrLeu 313
QY 778 -----CGACATGTGAGATCACTTCCCGCTCTCTG-----ATTCTG 813
Db 314 IlePheArgAsnLysLysGluIleThrGlnHisLeuLeuAsnPheAlaGluHisSerLeu 333
QY 814 GTGCTCTTCTCAGCCACTTTGAAA-----TTGAAGGCTGTG 849
Db 334 AlaPhePheSerLeuPheIleArgAlaPheAlaLeuValTrpHisTrpLeuAlaSerAla 353
QY 850 CCCTTCTAGTGTCTCAACTCTTCATCTTCATCTTGTGAGCCCTGGATTAAAGTTCTGAGA 909
Db 354 TyrPheIleValLeu-----PhePhePheSerLeuPheAspPro-----366
```

Alignment Scores:

Pred. No.:	0.697	Length:	521
Score:	104.50	Matches:	72
Percent Similarity:	37.77%	Conservative:	70
Best Local Similarity:	19.15%	Mismatches:	139
Query Match:	4.11%	Indels:	95
DB:	1	Gaps:	16

US-09-768-781-2 (1-1389) x YT25 CAEEL (1-521)

385	Qy	GCATGATTAAAGTACCTCACACTGTGGAGAAAGAGGAGGAGCCCTATATGTCAGC	444
168	Db	AlaAlaLeuAspPheLeuMetMetTrp-----TyrThrTrp	179
445	Qy	CTCACCCGAAAGAGATGCTAAATAGATGGCAGGAGGTGCTGATAGATGGGAGGTGGC	504
180	Db	AsnAspArgIysIleAsnTyrAspLeuAsnSerGlnTyrLeuLysValGluGlnHis	199
505	Qy	CACATCCATCCGGACCCCTGGCTATGCACACCGCAATGCCTACAAACGATGTGTGCACAGATCAAA	564
200	Db	HisSerLeuMetAlaValSerLeuAsn-----SerLeuLeuGln	212
565	Qy	GCCTTCTCGGCTCAGTGGCCCCAGCTGACCTATCACGCTCATGTGAGCCTGATCTCTCGCA	624
213	Db	LeuPheValThrSerIleTyrAlaIleSerMetPheValLeuAlaAsnMetSerMetThr	232
625	Qy	GAGGTCCCTCGGTAGAGTT---GTGCTAAATGGTATTTTCCCTGGTATCTGTCACTTAT	681
233	Db	AsnProGluLeuGlyAsnAlaAsnLeuLeuArgTrpPheTyrAlaPheLysValLeuTyr	252
682	Qy	GGGGCCACCCCTTTGGCAAT-----ATGCTGGCTATCCAGATCAAGTACCAT	726
253	Db	ArgArgSerValPheHisIleSerThrHisGlnLeuLeuPheMetSerIleValTyrAla	272
727	Qy	GACTAC---AAGATTCCGCTTGGGCCACCTAGAGTCCTCTGTCATCACCATCTGCGCGACA	783
273	Db	AspPheHisAlaIleSerTyrGlyPheLeuGln-----AlaIleSerLeuTyrArgSer	290
784	Qy	TTGGAGATCATCTCCCGCCTCTGATTCCTGGTGCTCTCTCAGACCACCTTGAATTGAAG	843
291	Db	ValValPheTyrAspAsp---ProValMetIleMetPheThrGly-----	304
844	Qy	GCTGTGCCCTTCCTAGTGTCTCAACTCTCTCGATCATCTCTTTGGACCCCTGGATTAAAGTTC	903
305	Db	-----TyrGluCysTyrAlaIleThr	311
904	Qy	TGGAGAAAGTGTCGC-----CAGATGCCCAATACATTGAG	939
312	Db	TrpThrValGlyAlaAlaLysMetLeuMetSerPheIleGlnThrAlaLeuThrIleGlu	331
940	Qy	AAAAACTTCAGCCGGGTGGCAGCTCTGGTG-----	969
332	Db	ArgIleIleAspArgIleIleProLeuValProLysLeuLysProPheLysArgGlnGly	351
970	Qy	GTCTGTGATTACGTACCATCTCTATGTGGCATCAACTCTCTGTGGTGGCTGCTTTG	1029
352	Db	LeuPheLeuAsnAlaPheAlaLeuThrAlaGlyIleSerThrThrIleTyrSerTyrSer	371
1030	Qy	CAG-----TTGAGGTTGGCA-----GAC	1047
372	Db	GluGlyProThrGlyTyrLysLeuAlaSerCysPheMetGlnLysAspIleProLeuAsp	391
1048	Qy	AGAGATCTCGTCGACAAAGGGCAGAACTGGGACATATGGGCTCGCATATAGTGTGAGG	1107
392	Db	ArgValLeuTyrThrLeuGlyMet-----TyrPheAlaLeuSerLeuLeuCysLeu	408
1108	Qy	TTGTGAGAGAATGTGATCATGCTCTTGGTTTTT-----AAGTTCTTTGGAGT	1155
409	Db	LeuAlaAsnLeuThrIleIlePheSerIlePheLysSerSerLysLysSerPheAsnLeu	428
1156	Qy	AAAGTGTACTGAAT-----TACTGTTCATCTCTGATGCC	1191
429	Db	LysIleArgPheAsnPheGlnGluValLysAsnSerSerPheAlaValSerIleLysSer	448

Qy	1192	TTGCAGCTCATTAATTCCTTATCTGATTTCCATTGGCTTCATGCTCCTTTTCTTCAGTAC	1251
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	449	ValPheGlnPheValAlaMetThrAlaTyValAlaSerSerLeuIleValIleTyTyr	468
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Qy	1352	TTGCATCCATTGGGCTCATCTCTCCACCCATAATAGTAGAGACTACCTCCATTGGTGTGTC	1311
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	469	ArgArgAspIleProHisGluTyPheHisAsnIle-----LeuCys	483
Qy	1312	TGTCACGACGACCTCGGACCAAGGTTTGAGAACTCAGAGCCACCCCTTT	1359
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	484	IleTyValSerLeuIlysLeuGlnHisLeuThrCysProProPhe	499

RESULT 6

ID	BAIL_HUMAN	STANDARD	PRT	1584 AA.
AC	014514;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Brain-specific angiogenesis inhibitor 1 precursor.			
GN	BAIL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98054121; PubMed=9393972;			
RA	Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,			
RA	Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;			
RT	"A novel brain-specific p53-target gene, BAIL, containing			
RT	thrombospondin type 1 repeats inhibits experimental angiogenesis."			
RL	OncoGene 15:2145-2150(1997).			
RL	[2]			
RP	INTERACTION WITH BAP1.			
RX	MEDLINE=98321173; PubMed=9647739;			
RA	Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,			
RA	Tokino T.;			
RT	"Cloning and characterization of BAI-associated protein 1: a PDZ			
RT	domain-containing protein that interacts with BAIL.";			
RL	Biochem. Biophys. Res. Commun. 247:597-604(1998).			
CC	-1- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN			
CC	BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53			
CC	SIGNAL IN SUPPRESSION OF GLOBLASTOMA. MAY FUNCTION IN CELL			
CC	ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.			
CC	-1- SUBUNIT: INTERACTS WITH BAP1.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE			
CC	CONCENTRATED AT CELL-CELL ADHESION SITES.			
CC	-1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN BRAIN. REDUCED ON			
CC	EXPRESSION IS OBSERVED IN SOME GLOBLASTOMA CELL LINES AND CANCER			
CC	TISSUES.			
CC	-1- INDUCTION: BY P53.			
CC	-1- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT			
CC	CORNEA INDUCED BY BFGF.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS			
CC	-1- SIMILARITY: CONTAINS 5 TSP TYPE-1 DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 1 GPS DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collab-			
CC	oratory agreement between the Swiss Institute of Bioinformatics and the EMBL outstations			
CC	the European Bioinformatics Institute. There are no restrictions			
CC	use by non-profit institutions as long as its content is in			
CC	modified and this statement is not removed. Usage by and for com-			
CC	panies requires a license agreement (See http://www.isb-sib.ch/submit			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB005297; BAA23647.1; -			
DR	GenBank; HGNC:943; BAIL.			
DR	MF; 602682; -			
DR	InterPro; IPR000832; GPCR_secretin.			
DR	InterPro; IPR00203; PKD_Cys_rich.			
DR	InterPro; IPR000884; TSP1.			

Qy	1078	GGACATATGCGCCTGACCTATATGCTGTGAGGTTGGTAGAGAATGTGATCATGCTCTTG---	1134
Db	366	---AsnLeuAlaLeuProThrIleAsnLeuValGlyGluLeuPheValValMetSer	384
Qy	1135	GTTTTAAAGTCTTTGGAGTGAAGTGTTACTGAATTACTGTCATCTCCTGATTCGCTTG	1194
Db	385	SerPheSerTrpSerAsnIleThrIleIle-----LeuMetGlyThr	398
Qy	1195	CAGCTCATTTATCTTATCTTGATTTTCATTTGGCTTCATGCTCTCCTTTCTTCAGTACTTG	1254
Db	399	AsnIleIleIleThrAlaLeuTyrsrLeu---TyrMetLeuIleThrThrGln-----	415
Qy	1255	CATCATTGGCGTCACTCTTCCACCAATATAGTAGAGACTACCTCCATGTGTCTGCTGT	1314
Db	416	-----ArgGlyLysTyThrHisHis-----	422
Qy	1315	CACCAGCACCTCGGACCAGGGTGTGAGACTCAGAGCCACCTTTTGAGACTGAA	1368
Db	423	-----IleAsnAsnIleLysProSerPheThrArgGlu	433
RESULT 9			
YN8I	YEAST		
ID	YN8I	STANDARD;	PRT; 551 AA.
AC	PS3730;		
DT	01-OCT-1996	(Rel. 34, Created)	
DT	01-OCT-1996	(Rel. 34, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	
DE	Hypotheical 62.7 kDa protein in SEC12-SSK2 intergenic region.		
GN	YN030W OR N3265.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
NCBI_TaxID	4932;		
OR	[1]		
RN	SEQUENCE FROM N.A.		
RA	Pohl T.M.;		
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).		
CC	-1- SIMILARITY: TO C.ELEGANS 2C5I3.5 AND S.POMBE SPBC1734.12C.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-		
CC	tion between the Swiss Institute of Bioinformatics and the EMBL Outstati-		
CC	on the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no		
CC	modified and this statement is not removed. Usage by and for commer-		
CC	entities requires a license agreement (See http://www.isb-sib.ch/annou		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; Z71645; CAA96310.1; -.		
DR	SGD; S0005313; YNR030W.		
KW	Hypotheical protein; Transmembrane.		
FT	TRANSMEM 3 23	POTENTIAL.	
FT	TRANSMEM 62 82	POTENTIAL.	
FT	TRANSMEM 90 110	POTENTIAL.	
FT	TRANSMEM 137 157	POTENTIAL.	
FT	TRANSMEM 179 199	POTENTIAL.	
FT	TRANSMEM 203 223	POTENTIAL.	
FT	TRANSMEM 228 248	POTENTIAL.	
FT	TRANSMEM 276 296	POTENTIAL.	
FT	TRANSMEM 304 324	POTENTIAL.	
FT	TRANSMEM 332 352	POTENTIAL.	
FT	TRANSMEM 366 386	POTENTIAL.	
FT	TRANSMEM 418 438	POTENTIAL.	
SQ	SEQUENCE 551 AA; 62672 MW; B08A94BBF260502F CRC64;		
Alignment Scores:			
Pred. No.:	1.82	Length:	551
Score:	99.50	Matches:	57
Percent Similarity:	32.66%	Conservative:	72
Best Local Similarity:	18.23%	Mismatches:	115
Query Match:	3.91%	Indels:	151
DB:	1	Gaps:	15

Alignment Scores:	
Pred. No.:	1.82
Score:	99.50
Percent Similarity:	32.66%
Best Local Similarity:	18.23%
Query Match:	3.91%
DB:	1
Length:	551
Matches:	72
Conservative:	57
Mismatches:	115
Indels:	151
Gaps:	15

US-09-768-781-2 (1-1389) x YN8I YEAST (1-551)

259	QY	TTTATGTTTTTCATCCATTATGGTCGACCTTGACCCCTCATTTTTTGTCCACAGAGACTGACCC	318
75	Db	TyrLeuTyrLeuSerSerLeuIleGlnThrSer-----	85
319	QY	AAAGATAAACCGCTATCATTTATTTATGTCATCTAATCCTCTTGGGACCTGTT-----	369
86	Db	-----ArgProThrSerIleAspValGlnLeuValValArgGlyIleValGlyLeuThr	103
370	QY	-----ATCAGATCTTTGGAGGCCATGATTAAAGTACCTCCTCAC	405
104	Db	AsnGlyLeuSerPheIleTyrLeuLysAsnCysLeuGlnAspMetPheAspGluIleThr	123
406	QY	CTGTGGAAGAAAGAGGAGGAGGAGGCC-----TATGTC-----	441
124	Db	GluLysLysGluAsnGluAspLysAspIleTyrIleTyrAspSerAlaGlyThr	143
442	QY	-----AGCCTC-----	447
144	Db	TrpPheLeuLeuPheLeuIleGlySerPheHisLeuMetPheTyrSerThrArgThrLeu	163
448	QY	ACCGAAAGAGATGCTAATAGATGCGGAGGAGGTGCTAGTAGAATGGGAGTGGGCCAC	507
164	Db	ProAsnPheValMetThrLeuProLeuThrAsnValAlaLeuGlyTrp-----	179
508	QY	TCCATCCGACCTCGCTATGCACCGCAATGCTACAAACGTATGTCACAGATCCAAGCC	567
180	Db	-----ValLeuLeuGlyArgTyrAsnAla-----AlaIle	189
568	QY	TTCTGGGCTCAGTCCCGCCAGCTCATAGCTCTATGTGAGCGTGAATCTGCAGAG	627
190	Db	PheLeuSerAlaLeuValAlaIleValPheArgLeuGluValSerAlaLeuSerAlaGly	209
628	QY	GTTCCTCCCTGGTAGATTGTGCTAATGGTATTTTCCTGGTATCTGTC-----	675
210	Db	IleAlaLeuPheSerValIlePheLysLysIleSerLeuPheAspAlaIleLysPheGly	229
676	QY	ACCTATGGGGCCACCTTTGGCAATATGTTGGCTATTCAGATCAAGTACGATGACTAC--	732
230	Db	IlePheGlyLeuGlyLeuGlySer-----AlaIleSerIleThrValAspSerTyrPhe	247
732	QY	-----	732
248	Db	TrpGlnGluTrpCysLeuProGluValAspGlyPheLeuPheAsnValValaGlyTyr	267
733	QY	-----NAGATTCCGCTTGGCCCATAGAGTCCTCTGCATCACCATCTGGCGGACATTG	786
268	Db	AlaSerIleTrpGlyValGluProValThrAlaTyrPheThrHisTyrLeuArgMetMet	287
787	QY	GAGATCACTTCCCGCTCTGATCTCGGTGCTCTCTCA-----GCCACTTTGAAA	837
288	Db	PheMetProThrValLeuLeuLeuAsnTyrPheGlyTyrLysLeuAlaProAlaLys	307
838	QY	TTGAGGCTGTGCCCCTCTCCTAGTGTCCAACTCTCTGATCATCTCTC-----TTTGAGCCCTCG	894
308	Db	LeuLysIleValSerLeuAlaSerLeuPheHisIleIleValLeuSerPheGlnProHis	327
895	QY	ATTAAAGTTCTGGAGAAGTGGTGGCCAGATGCCCAATAACATTGAGAAAACTTCAGCCGG	954
328	Db	---LysGluTrpArg-----	331
955	QY	GTCCGCACTCTGGTGGTCCTGATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCT	1014
332	Db	-----PheIleIleTyrAlaValProSerIleMetLeuLeuGlyAlaThrGlyAla	348
1015	QY	TGCTGTGTGAGCTTTGTCAGTTTGGAGTTGGGCAGACAGAGATCTCGTGCACAAAGGCGAGAAC	1074
348	Db	-----	348
1075	QY	TGGGGACATATGGCCCTGCATATAGTGTGAGTTGGTAGAATGTGATCATGGTCTTG	1134
349	Db	---AlaHisLeuTrpGluLeuAsnMetLysValValValIleThrAsnValLeu-----	364


```

FT TRANSMEM 429 447 POTENTIAL.
FT DISULFID 153 167 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 36 38 LFO -> EWK (IN REF. 3).
SQ SEQUENCE 458 AA; 52728 MW; D6E919E53CBD21F8 CRC64;

Alignment Scores:
Pred. No.: 1-94 Length: 458
Score: 99.00 Matches: 45
Percent Similarity: 45.71% Conservative: 35
Best Local Similarity: 25.71% Mismatches: 63
Query Match: 3.89% Indels: 8
DB: 1 Gaps: 8

US-09-768-781-2 (1-1389) x ACHO_HUMAN (1-458)
QY 256 TTCCTTTATGTTTCATCCATTATGGTCAGTGCACCTCATTT---TTTGTC--- 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 TyrLeuLeuPheileMetilePheValThrLeuSerileleValThrValPheValle 318
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 -----CACAGATCTAGCAAGATAAACCGTATCATATTATTATGATCTCAATC 354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 319 AsnValHisArgSerSerThrTyrHisProMetAlaProTrpValLysArgLeu 338
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 355 CTCCTGGGA-----CTGTTATCAGATGTTGGAGGCCATGATTAACTACCTCACACTG 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 PheLeuGlnLysLeuProLysLeuLysCysMetLysAspHisValAspArgTyrSerSer 358
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 409 TGAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 ProGluLysGluGlu---SerGlnProValValLysGlyLysValLeuGluLysLysLys 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 454 AGAAGATGCTAATAGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 507
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 378 GlnLysGlnLysSerAspGlyGluLysValLeuPheLeuGluLysAlaAlaAsp 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 508 TCATCCGACCGCTGCTATGACCGCAATGCTTACAAAGTATGTCACAGATCCAGGCC 567
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 SerileArgTyrileSerArgHisValLysLysGluHisPheileSerGlnValGln 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 568 TTCCTGGGCTCAGTGCAGCCAGCTGCATCAGCTCTATGAGCCTGATCTCTGCAGAG 627
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 AspTrpLysPheValAlaGlnVal----- 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 628 GTTCCCGCTGGTAGAGTGTGCTAATGTTATTTCCCTGCTATCTGTCACTATGGGCGC 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 426 -----LeuAspArgilePheLeuTrpLeuPheLeuValSerValThr---GlySer 442
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 688 ACCCTTTCGAATATGTTGCTATCCAGATCAAGTACGATGACTAC 732
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 443 ValLeuPhePheThrProAlaLeuLysMetTrpLeuHisSerTyr 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
VPHL_SCHPO STANDARD; PRT; 805 AA.
AC 013742;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable vacuolar ATP synthase 91 kDa subunit (Vacuolar ATPase 91 kDa subunit).
GN SPAC1688.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;

```

```

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Spouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Horsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: REQUIRED FOR ASSEMBLY AND ACTIVITY OF THE VACUOLAR
CC ATPASE. POTENTIAL ROLE IN DIFFERENTIAL TARGETING AND REGULATION OF
CC THE ENZYME FOR A SPECIFIC ORGANELLE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. VACUOLAR
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE V-ATPASE 116 kDa SUBUNIT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z98529; CAB11035.1; -.
CC InterPro; IPR002490; V_ATPase_sub116.
CC Pfam; PF01496; V_ATPase_sub_a7_1.
KW Hydrogen ion transport; Transmembrane; Glycoprotein.
FT DOMAIN 1 404 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 405 429 1 (POTENTIAL).
FT DOMAIN 430 451 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 452 472 2 (POTENTIAL).
FT DOMAIN 473 526 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 527 547 3 (POTENTIAL).
FT DOMAIN 548 558 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 559 579 4 (POTENTIAL).
FT DOMAIN 580 590 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 591 611 5 (POTENTIAL).
FT DOMAIN 612 624 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 625 645 6 (POTENTIAL).
FT DOMAIN 646 685 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 686 706 7 (POTENTIAL).
FT DOMAIN 707 745 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 746 766 8 (POTENTIAL).
FT DOMAIN 767 805 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 805 AA; 91226 MW; DE92CCBACC66C6CFC CRC64;

Alignment Scores:
Pred. No.: 2-58 Length: 805
Score: 98.00 Matches: 97
Percent Similarity: 37.47% Conservative: 78
Best Local Similarity: 20.77% Mismatches: 174
Query Match: 3.85% Indels: 118

```

[illegible]

Db	664	AnValAaspGluAlaAaspAlaValAaspSerGlnGluLeuGlnAlaGluPro	683
QY	952	---CGGTCGGCAGCTCTGGTGC	972
Db	684	PheGluLeuGlyGluValValIleHisGlnValIleHisThrIleGluPheCysLeuGly	703
QY	973	CTGATTTTCAGTCACCATCTCTATGCTGGCATCAACTTCTCTGTGTCGTCAGCTTTCAG	1032
Db	704	CysValSerHisThrAlaSerTyrLeuArgLeu	716
QY	1033	TTGAGTTGGCAGACAGATCTCTGCACAAAGGCGACACTGGGACATATGGCCCTG	1092
Db	717	LeuSerLeuAlaHisAsnGlnLeu	733
QY	1093	CACATATAGTGTGAGG	1143
Db	734	AlaAsnGlyPheArgMetThrClyIleValGlySerIlePheValIleLeuPheGly	753
QY	1144	TTCTTTGGAGTGAAGTGTACTGAATTAAGTCTGATCTGCTGATTCGCTGAGCTCATT	1203
Db	754	PheTyrPheIleAlaThr	769
QY	1204	ATTGCTTATCTGATTTCCATT	1224
Db	770	SerAlaMetLeuHisSerLeu	776
RESULT	13		
MS5R	RAT		
ID	MC5R_RAT	STANDARD;	PRT; 325 AA.
AC	P35345;		
DT	01-JUN-1994	(Rel. 29, Created)	
DT	01-JUN-1994	(Rel. 29, Last sequence update)	
DT	01-NOV-1995	(Rel. 32, Last annotation update)	
DE	Melanocortin-5 receptor (MC5-R).		
GN	MC5R.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Striatum;		
RC	MEDLINE=94234987; PubMed=8179577;		
RA	Griffin N., Mignon V., Facchinetti P., Diaz J., Schwartz J.C.,		
RA	Sokoloff P.;		
RT	"Molecular cloning and characterization of the rat fifth melanocortin		
RL	receptor.";		
RL	Biochem. Biophys. Res. Commun. 200:1007-1014(1994).		
CC	!- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA AND GAMMA) AND ACTH. THE		
CC	ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE		
CC	ADENYLATE CYCLASE. THIS RECEPTOR IS A POSSIBLE MEDIATOR OF THE		
CC	IMMUNOMODULATION PROPERTIES OF MELANOCORTINS.		
CC	!- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	!- TISSUE SPECIFICITY: VERY LOW EXPRESSION LEVELS IS DETECTED IN		
CC	BRAIN, WHILE HIGH LEVELS ARE FOUND IN ADRENALS, STOMACH, LUNG AND		
CC	SPLEEN.		
CC	!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collabora		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	use by non-profit institutions as long as its content is in no		
CC	modified and this statement is not removed. Usage by and for commerc		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announ		
CC	or send an email to license@isb-sib.ch)		
CC	EMBL; L27081; AAA41577.1; ..		
DR	PIR; JC2193; JC2193.		
DR	InterPro; IPR000276; GPCR_Rhodopsn.		
DR	Pfam; PF00001; 7tm_1; 1.		
DR	PROSITE; PS00237; G-PROTEIN RECF Fl_1; 1.		
DR	PROSITE; PS0262; G-PROTEIN RECF Fl_2; 1.		

CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -I- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; X54860; CAA38643.1; -.
DR ENBL; U03843; AAB93444.1; -.
CC ENBL; X66484; CAA47115.1; -.
DR ENBL; X12939; -I-NOT_ANNOTATED_CDS.
DR PIR; S09172; S09172.
DR PIR; S28109; S28109.
DR InterPro; IPR001750; Oxidored q1.
DR Pfam; PF00361; oxidored q1; 1.
CC KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
FT CONFLICT 366 366 A -> R (IN REF. 3).
SQ SEQUENCE 382 AA; 42382 MW; 358658E8EDCFCAE9 CRC64;

Alignment Scores:
Pred. No.: 2.74 Length: 382
Score: 97.00 Matches: 56
Percent Similarity: 38.93% Conservative: 53
Best Local Similarity: 20.00% Mismatches: 83
Query Match: 3.81% Indels: 88
DB: 1 Gaps: 12

US-09-768-781-2 (1-1389) x NU2M_CHLRE (1-382)

Qy 613 CTGATCTCTGCAGAGGTTCCCTGGGTAGAGTTGTGCTAATGGTATTTTCCTGGTATCT 672
Db 1 MetileGluLeuAspLeuCysPheGlyLeuLeuLeuLeuLeuPheGlyLeuLeuSer 20

Qy 673 GTACCTATGCGGCCACCCCTTGGCAATATGTTGGCTTCAGATCAAGTAGTAGACTAC 732
Db 21 LeuArgAsnGlyHisVal-----SerLeuAlaHisIleArgPheIleCysGlnCysTrp 38

Qy 733 AAGATTCGCTTGGGCCACTAGAGTC-----CTGTCATCACCATCTGCGCG 780
Db 39 LeuValMetIleThrProLeuGluValGlnAspPheAlaLeuCysIle----- 54

Qy 781 ACATTGGAGATCATCTCCGCTCTGATTCGTGTGCTCTCTCAGGCACCTTTGAAATTG 840
Db 55 -----LeuThrValValLeuLeuGlnSerPheHisSerPhe 66

Qy 841 AAGCTGTGCCCTTCCTAGTGTCAACTTCTCTG----- 873
Db 67 GluAlaLeuLeuPheLeuLeuLeuAlaTyrlleGlyGlnLeuTyMetMetHisSerCys 86

Qy 874 ---ATCATCTCTTTAGCCCTCGGATTAAGTTCTGGAGAAAGTGTGCGCCAG----- 921
Db 87 AsnLeuValSerPheTyValCysLeuGlu-----AlaGlnThrLeuCys 101

Qy 922 -----ATCCCAATACATTGAGAAAAAATCTCAGCCGGTCCGC 960
Db 102 ValValValLeuCysGlyLeuLeuAlaArgGlyAlaSerThrSerPheSerValGluAla 121

Qy 961 ACTCTGGTGGTCTGATTTTCAGTCACCATCTCTATGCTGGCATCAACTTCTCTGCTGG 1020
Db 122 AlaLeuLysPheLeuLeu---LeuSerAlaMetValSerGlyMetAlaLeuPheTrpPhe 140

Qy 1021 TCAGCTTTGAGGTTGGAGTGGCCAGACAGAGATCTCGTCACAAAAGGCGCAACTGGGGA 1080
Db 141 SerAlaMetTyGlnAArgThrGlySerLeuAspMetValGlyGlnGluThrPheTrp--- 159

Qy 1081 CATATGGCCCTGCACATAATAGTGTGAGGTGGTAGAGAAATGTGATCATGGTCTTGTTT 1140
Db -----


```
QY 184 GCATCTCTTGTATCATGGTTAGAAATCTATCGAAGAAATAGTGAAGAACTTACTGGATGACA 243
Db 58 ThrPheSerValLeuValPheSerThrTyTyTyLeuAsnSerGlu----- 72
QY 244 TACACCTTTCTTCTTATGTTTTCATCCATTATGGTCCAGTTGACCCCTCAATTTTGTGTC 303
Db 73 LeuAsnPheAsnTyTyTyTyPhe-----ValLeuLeuLeuPheVal 86
QY 304 CACAGAGATCTAGCCAAAGATAAACCCTATCATATTATGCTTAATCTCTTGGGA 363
Db 87 GlySerMetPheSerLeuAsnPheSerAsnSerIlePheThrMetLeuLeu-Leu-SerTrpAs 106
QY 364 CTTGTTATCAGA----- 380
Db 106 pLeuLeuGlyIleSerSerPhePheLeuValLeuPheTyTyAsnAsnTrpAspSerCysSe 126
QY 381 GGAGGCCATGATTAAAGTACCTCACACTGTGGAAGAAAGAGGAGGAGGAGCCCTATGT 440
Db 126 rGlyAlaMetAsnThrAlaLeuThr-----AsnArgLeuGlyAspTyTyPh 141
QY 441 CAGCCTCACCCGAAAGAGATGCTAATAGATGGCGAGGAGTGTCTGATAGAAATGGGAGGT 500
Db 141 eMetPheValPhePheGlyLeuSerValPheSerGlyTyTyTyPheLeuSerPheSerMe 161
QY 501 GGCGCCACTCCATCCGACCTCGGTATGCACCGCAATGCCACAAACGTATGTCCACAGAT 560
Db 161 tPheSerSerTyMetSerLeuLeuLeuLeuLeuLeuAlaPheThrLySerAlaGlnPh 181
QY 561 CCAAGCCTTCTGGGCTCAGTGGCCAGCTGACCTATCAGCTCTATGTGAGCCGTGATCTC 620
Db 181 e---ProPheSerSerTrpLeuProLys-----AlaMetSe 192
QY 621 TGCAGAGGTTCCCTCGGTAGATGTGTAAATGCTATTTTCCCTGCTATCTGTCACTA 680
Db 192 rAlaProThrPro-----ValSerSerLeuValHisSerSerThrLeuValThrAl 209
QY 681 TGGGGCCACCCCTTTGC-----AATATGTTGGCTATCCAGATCAAGTACGATGACTA 731
Db 209 aGlyLeuLeuLeuLeuMetAsnPheAsnAsnLeuValMetGlnLysAspPheIleSerPh 229
QY 732 CAAGATTCGGCCTTTGG----- 747
Db 229 eValLeuIleIleGlyLeuPheThrMetPhePheSerSerLeuAlaSerLeuValGluGl 249
QY 748 -----CCACTAGAAGTCTC-----TGCATCAC 770
Db 249 uAspLeuLysValValAlaLeuSerThrLeuSerGlnMetGlyPheSerMetValTh 269
QY 771 CATCTGGCGGACATTTGGAGATCACTTCC-----CGCCTCCTGATTTCTGTGCTCTTCTC 824
Db 269 rLeuGlyLeuGlyLeuSerPheIleSerPheIleHisLeuValSerHisAlaLeuPheLy 289
QY 825 AGCCACTTGAATTTGAAGCT----- 846
Db 289 sSerCysLeuPheMetGlnValGlyTyTyIleIleHisCysSerPheGlyGlnGlnAspGl 309
QY 847 -----GTGCCC-----TTCTAGTGTCTCACTTCTGTATCAT 878
Db 309 yArgAsnTySerAsnAsnGlyAsnLeuProAsnPheIleGlnLeuGlnMetLeuValTh 329
QY 879 CTTCTTTGAGCCCTGGATTAAAGTTCTGGAGAAGTGGTGGCCAGATGCCCAATACATT-- 936
Db 329 rLeuPheCysLeuCysGlyLeuIlePheSerSerGlyAlaValSerLysAspPheIleLe 349
QY 937 -----GAGAAAAAATTCACCGCGGTTCGGCACTCTGGTGGTCTCTGATTTTCAGT 983
Db 349 uGluLeuPhePheSerAsnAsnTyMetMetPhePheSerLeuMetPhePheValSerVa 369
QY 984 CACCATCTCTATGCTGGCATCAACTTCTCTTGTGTGTGCTGAGTTGAGGTTGGC 1043
Db 369 lPheLeuThrPhe---GlyTySerPheArgLeuTrpLysSerPhePheLeuSerPhe-- 387
QY 1044 AGACAGAGATCTCGTGACAAAGGGCAGAACTGGGGACATATGGGCTGCACTATAGTGT 1103
```

```
Db 388 -----AsnLysValMetAsn-----HisTySerSe 396
QY 1104 GAGTTTGGTAGAGAAATGTGATCATGGTCTTGGTTTAAAGTTCTTTGGAGTGAAAGTGT 1163
Db 396 rThrValPheMetAsn---PheLeuSerLeuValLeuValIlePheSerIleSerPheLe 415
QY 1164 ACTGAATTACTGTCTATTCCTTGGATTGGCTTGCAGCTCATATTGCTTATCTGATTCCAT 1223
Db 415 uTrpTrpMetAsnPheAsnLeuLeuAsnIleProSerLeuPheLeuTyTyValAspPhePh 435
QY 1224 TGGC-----TTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTCGCGCTC 1268
Db 435 eGlyProLeuValPheLeuPheMetMetIlePheLeuSerPheLeu---IleLeuLysMe 454
QY 1269 ACTCTTCACCCAT-----AATGTAGTAGACTACCTC 1299
Db 454 tLeuPheLysGluLeuMetTyTyLysPheLeuValAspTyTyLeu 468
```

Search completed: April 1, 2003, 08:54:20
Job time : 36.5 secs